

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2001, 14:39:33 ; Search time 19.61 Seconds
(without alignments)
4262.284 Million cell updates/sec

Title: US-09-522-753-11

Perfect score: 12643
Sequence: 1 MSSSGYPPNQAGAFSTQSR.....EREAPLLSAQYETLSDSD 2440

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12643	100.0	2440	1 NCRL_HUMAN	075376 homo sapien
2	11536.5	91.2	2453	1 NCRL_MOUSE	060974 mus musculus
3	4147.5	32.8	2517	1 NCRL_HUMAN	095618 h nuclear r
4	4055.5	32.1	2472	1 NCRL_MOUSE	094042 mus musculus
5	2493.5	19.7	533	1 NCRL_RAT	Q9W055 rattus norv
6	480	3.8	2468	1 MAPB_HUMAN	P46821 homo sapien
7	465.5	3.7	2774	1 MAPB_RAT	P34926 rattus norv
8	447	3.5	2464	1 MAPB_MOUSE	P14873 mus musculus
9	437.5	3.5	2805	1 MAPA_HUMAN	P78559 homo sapien
10	437	3.5	2845	1 APC_MOUSE	Q61315 mus musculus
11	411.5	3.3	3924	1 ANK2_HUMAN	Q01484 homo sapien
12	406	3.2	2842	1 APC_RAT	P70478 rattus norv
13	404	3.2	3256	1 KIF7_HUMAN	P46013 homo sapien
14	390	3.1	1226	1 YCS3_YEAST	P25357 saccharomyc
15	374.5	3.0	3562	1 PCV_CHICK	Q09053 gallus gall
16	365	2.9	2843	1 APC_HUMAN	P25054 homo sapien
17	360	2.8	2349	1 TPR_HUMAN	P12270 homo sapien
18	353	2.8	2142	1 BAT2_HUMAN	P48634 homo sapien
19	350	2.8	661	1 YB3_SCHPO	Q10369 schizosacch
20	348.5	2.8	1018	1 YK26_CAEEL	P34333 caenorhabdi
21	343.5	2.7	1411	1 TCOF_HUMAN	Q13428 homo sapien
22	343	2.7	3866	1 HRX_HUMAN	P55200 mus musculus
23	336	2.7	2688	1 ZEP1_MOUSE	Q03172 mus musculus
24	332.5	2.6	1861	1 MAP2_RAT	P15146 rattus norv
25	326	2.6	1781	1 AKAC_HUMAN	Q02952 homo sapien
26	321.5	2.5	3969	1 HRX_HUMAN	Q03164 homo sapien
27	317	2.5	4687	1 PLE1_RAT	P30427 rattus norv
28	314.5	2.5	3358	1 PCV_MOUSE	Q62059 mus musculus
29	312	2.5	1902	1 SMF1_HUMAN	Q14497 homo sapien
30	310.5	2.5	2004	1 MO2_HUMAN	Q92794 homo sapien
31	309.5	2.4	2441	1 CBP_MOUSE	P45481 mus musculus
32	308	2.4	2442	1 CBP_HUMAN	Q92793 homo sapien
33	303	2.4	3396	1 PCV_HUMAN	P13611 homo sapien

34	300.5	2.4	2090	1 HFC1_MESAU	P51611 mesocricetu
35	299.5	2.4	2715	1 TRX2_HUMAN	Q9U0N6 homo sapien
36	298.5	2.4	771	1 CALD_CHICK	P12957 gallus gall
37	298	2.4	2194	1 SC16_YEAST	P48415 saccharomyc
38	296.5	2.3	1828	1 NFM_MOUSE	P20357 mus musculus
39	295.5	2.3	848	1 NFM_HUMAN	P08553 mus musculus
40	295.5	2.3	2035	1 HFC1_HUMAN	P51610 homo sapien
41	295	2.3	2717	1 ZEP1_HUMAN	P15822 homo sapien
42	294	2.3	915	1 NFM_HUMAN	P07197 homo sapien
43	291	2.3	1827	1 MAP2_HUMAN	P11137 homo sapien
44	289	2.3	1637	1 MRSP_STAAU	P80544 staphylococ
45	284.5	2.3	1505	1 CDP_HUMAN	P39880 homo sapien

ALIGNMENTS

RESULT	1
NCRL_HUMAN	
ID NCRL_HUMAN	STANDARD; PRT: 2440 AA.
AC 075376; Q9UPV5; Q9U018;	
DT 01-OCT-2000 (Rel. 40, Created)	
DT 01-OCT-2000 (Rel. 40, Last sequence update)	
DT 01-OCT-2000 (Rel. 40, Last annotation update)	
DE NUCLEAR RECEPTOR CO-REPRESSOR 1 (N-COR) (N-COR).	
GN NCOR1 OR KIN1047.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX NCBI_TaxID=9606;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC TISSUE=Brain;	
RX MEDLINE=98393736; PubMed=9724795;	
RA Wang J., Hoshino T., Redner R.L., Kajigaya S., Liu J.M.;	
RT "ETO, fusion partner in t(8;21) acute myeloid leukemia, represses	
RT transcription by interaction with the human N-COR/msin3/HDAC1	
RT complex.";	
RL Proc. Natl. Acad. Sci. U.S.A. 95:10860-10865(1998).	
RN [2]	
RP SEQUENCE OF 782-2440 FROM N.A.	
RC TISSUE=Brain;	
RX MEDLINE=99397452; PubMed=10470851;	
RA Kikuno R., Nagase T., Ishikawa K.-I., Hirosewa M., Miyajima N.,	
Tanaka A., Kotani H., Nomura N., Ohara O.;	
RT "Prediction of the coding sequences of unidentified human genes. XIV.	
RT The complete sequences of 100 new cDNA clones from brain which code	
RT for large proteins in vitro.";	
RL DNA Res. 6:197-205(1999).	
RN [3]	
RP SEQUENCE OF 974-2440 FROM N.A.	
RX MEDLINE=99375328; PubMed=10444336;	
RA Nagaya T., Chen K.-S., Fujieda M., Ohmori S., Richer J.K.,	
Horwitz K.B., Lupski J.R., Seo H.;	
RT "Localization of the human nuclear receptor co-repressor (hn-Cor) gene	
RT between the CMT1A and the SMS critical regions of chromosome	
RT 17p11.2.";	
RL Genomics 59:339-341(1999).	
CC -1- FUNCTION: MEDIATES THE TRANSCRIPTIONAL REPRESSION ACTIVITY OF SOME	
CC NUCLEAR RECEPTORS BY PROMOTING CHROMATIN CONDENSATION, THUS	
CC PREVENTING ACCESS OF THE BASAL TRANSCRIPTION.	
CC -1- SUBUNIT: FORMS A LARGE COREPRESSOR COMPLEX THAT CONTAINS SIN3A/B	
CC AND HISTONE DEACETYLASES HDAC1 AND HDAC2. THIS COMPLEX ASSOCIATES	
CC WITH THE THYROID (TR) AND THE RETINOIC ACID RECEPTORS (RAR) IN THE	
CC ABSENCE OF LIGAND.	
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).	
CC -1- DOMAIN: THE N-TERMINAL REGION CONTAINS REPRESSION FUNCTIONS THAT	
CC ARE DIVIDED INTO THREE INDEPENDANT REPRESSION DOMAINS (RD1, RD2	
CC AND RD3). THE C-TERMINAL REGION CONTAINS THE NUCLEAR RECEPTOR-	
CC INTERACTING DOMAINS THAT ARE DIVIDED IN TWO SEPARATE INTERACTION	
CC DOMAINS (ID1 AND ID2).	
CC -1- DOMAIN: THE TWO INTERACTION DOMAINS (ID) CONTAIN A CONSERVED	
CC SEQUENCE REFERRED TO AS THE CORNR BOX. THIS MOTIF IS REQUIRED AND	

CC SUFFICIENT TO PERMIT BINDING TO UNLIGANDED TR AND RAS. SEQUENCES
CC FLANKING THE CORNR BOX DETERMINE NUCLEAR HORMONE RECEPTOR
CC SPECIFICITY.
CC -1- SIMILARITY: CONTAINS 1 SANT DOMAIN (COMPOSED OF SANT-A+SANT-B).
CC -1- SIMILARITY: CONTAINS 2 CORNR BOX.
CC -1- SIMILARITY: BELONGS TO THE N-COR NUCLEAR RECEPTOR COREPRESSORS
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC
CC EMBL: AF044209; AAC33550.1; -
CC EMBL: AB028970; BAA82999.1; -
CC EMBL: AB019524; BAA75814.1; -
CC MIM: 600849; -
CC InterPro: IPR001005; -
CC Pfam: PF00249; mvb_DNA-binding; 2.
CC PROSITE: P550090; MYB_3; 1.
CC Nuclear protein; Transcription regulation; DNA-binding; Repressor;
CC Coiled coil. 174 216 COILED COIL (POTENTIAL).
CC DOMAIN 234 312 INTERACTION WITH SIN3A/B.
CC DOMAIN 259 328 COILED COIL (POTENTIAL).
CC FT DNA_BIND 437 482 SANT-A (POTENTIAL).
CC FT DNA_BIND 625 670 SANT-B (POTENTIAL).
CC FT DOMAIN 501 557 COILED COIL (POTENTIAL).
CC FT DOMAIN 607 617 PRO-RICH.
CC FT DOMAIN 988 1816 INTERACTION WITH ETO.
CC FT DOMAIN 2055 2059 CORNR BOX OF ID1.
CC FT DOMAIN 2263 2267 CORNR BOX OF ID2.
CC FT DOMAIN 38 64 POLY-GLN.
CC FT DOMAIN 593 603 POLY-ALA.
CC FT DOMAIN 1032 1035 POLY-PRO.
CC FT DOMAIN 1707 1712 POLY-ALA.
CC FT DOMAIN 1952 1963 POLY-SER.
CC FT CONFLICT 1014 1014 L -> V (IN REF. 2).
CC FT CONFLICT 1508 1509 PP -> SS (IN REF. 2).
CC FT CONFLICT 1561 1561 W -> R (IN REF. 2).
CC FT CONFLICT 1567 1567 Q -> H (IN REF. 2).
CC FT SEQUENCE 2440 AA; 270263 MW; 60A4D7964D00EDAB CRC64;

Query Match 100.0%; Score 12643; DB 1; Length 2440;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSSSGYPPNQGAFSTEQSRYPHVSQVYTFNTRHQEFAYVDYRSSHLEVSQASQLLQ00 60
Db 1 MSSSGYPPNQGAFSTEQSRYPHVSQVYTFNTRHQEFAYVDYRSSHLEVSQASQLLQ00 60
Qy 61 QOOLRRRPSLLSEFPSPGSDRQERRTSYEPFHPGSPVDHDSLESKRPRLEQVSDSHFQ 120
Db 61 QOOLRRRPSLLSEFPSPGSDRQERRTSYEPFHPGSPVDHDSLESKRPRLEQVSDSHFQ 120
Qy 121 RVSAANVLPLVPLPGLRASADAKDPARGCKHEAPSSPISSGCGDDONASPSKLSKEE 180
Db 121 RVSAANVLPLVPLPGLRASADAKDPARGCKHEAPSSPISSGCGDDONASPSKLSKEE 180
Qy 181 LIQSHMDVRDREIAKYEQOILKKKQOOLEEAAKPEPEKPSPPVPEQKHRSIVQIIY 240
Db 181 LIQSHMDVRDREIAKYEQOILKKKQOOLEEAAKPEPEKPSPPVPEQKHRSIVQIIY 240
Qy 241 DENRKAEEAHKIFEGLPKVELPLYNQSDTKYHENIKTNQVMKKLILFFKRNHAR 300
Db 241 DENRKAEEAHKIFEGLPKVELPLYNQSDTKYHENIKTNQVMKKLILFFKRNHAR 300
Qy 301 KQREQKICORYDOLMEAEKKVDRIENNPRAKESKTRTYEYKQPPETRKOREQOERQ 360
Db 301 KQREQKICORYDOLMEAEKKVDRIENNPRAKESKTRTYEYKQPPETRKOREQOERQ 360

Db 301 KQREQKICORYDOLMEAEKKVDRIENNPRAKESKTRTYEYKQPPETRKOREQOERQ 360
Qy 361 RVQQRAGLSATIAHSEHEISEIDGLSEONNEKQMRQLSVIPPMHDAEQRRVKFINM 420
Db 361 RVQQRAGLSATIAHSEHEISEIDGLSEONNEKQMRQLSVIPPMHDAEQRRVKFINM 420
Qy 421 NGLMEDPMKVKYKDRQFNKVTWDEHKEIFKQFIQHPKNFGLIASYLERKSVPCVLYYL 480
Db 421 NGLMEDPMKVKYKDRQFNKVTWDEHKEIFKQFIQHPKNFGLIASYLERKSVPCVLYYL 480
Qy 481 TKKNENYKALVRNNGKRGNGQIARPSQEEKVEEEDKAETKKEEKEDEEKEDE 540
Db 481 TKKNENYKALVRNNGKRGNGQIARPSQEEKVEEEDKAETKKEEKEDEEKEDE 540
Qy 541 KEDSKENTKEKIDGTAEETEEDEQATPRGRTANSQGRKGRITRSMTNNEAASAAA 600
Db 541 KEDSKENTKEKIDGTAEETEEDEQATPRGRTANSQGRKGRITRSMTNNEAASAAA 600
Qy 601 AAATEEPPLPPPPPEPISTEPVETSRWTEEMEVAKKGLVEHGRNNAIAKAVGCTKSEA 660
Db 601 AAATEEPPLPPPPPEPISTEPVETSRWTEEMEVAKKGLVEHGRNNAIAKAVGCTKSEA 660
Qy 661 OCKNENYKALVRNNGKRGNGQIARPSQEEKVEEEDKAETKKEEKEDEEKEDE 720
Db 661 OCKNENYKALVRNNGKRGNGQIARPSQEEKVEEEDKAETKKEEKEDEEKEDE 720
Qy 721 ENPEDESEAVKPSDSDPENATSRGNTPEPAVELEPTTETAPSTSPSLAVPSTKPADESV 780
Db 721 ENPEDESEAVKPSDSDPENATSRGNTPEPAVELEPTTETAPSTSPSLAVPSTKPADESV 780
Qy 781 ETQVNDISIAETAEDMDVQDQEHSAEGSCVCDPPATKADSDVEVYRPNHASKVGGDN 840
Db 781 ETQVNDISIAETAEDMDVQDQEHSAEGSCVCDPPATKADSDVEVYRPNHASKVGGDN 840
Qy 841 TREDLDRASEKVEPRDEDLVVAQOINQAQPEPOSNDSSATCSADEDDVGEPEQRNPF 900
Db 841 TREDLDRASEKVEPRDEDLVVAQOINQAQPEPOSNDSSATCSADEDDVGEPEQRNPF 900
Qy 901 MDSKPSLLNPTGSIIVSSPLAPNLDLPOLQHRAAVIPPMVSCPTPCNIPITGTPVSGYALY 960
Db 901 MDSKPSLLNPTGSIIVSSPLAPNLDLPOLQHRAAVIPPMVSCPTPCNIPITGTPVSGYALY 960
Qy 961 ORHKAHESALLEQRORQOQIDECRSSTSCGTSKSPNREWEVLQAPHQLINTLPE 1020
Db 961 ORHKAHESALLEQRORQOQIDECRSSTSCGTSKSPNREWEVLQAPHQLINTLPE 1020
Qy 1021 GYRLPTTRTPRPPPLIPSSKTTVASEKPSFIMGSSISQGTPTGTYLTSHNQASYTQETPK 1080
Db 1021 GYRLPTTRTPRPPPLIPSSKTTVASEKPSFIMGSSISQGTPTGTYLTSHNQASYTQETPK 1080
Qy 1081 PSVGSISLGLPQOESAKSATLPYIKOEEFSPRSONSQPEGLVRAQHGCVVGTAGATQ 1140
Db 1081 PSVGSISLGLPQOESAKSATLPYIKOEEFSPRSONSQPEGLVRAQHGCVVGTAGATQ 1140
Qy 1141 EGSITRGTPSKISVESIPSLRSGSITQGTALPQGTPTALVKGSIIRMPEDSSPERG 1200
Db 1141 EGSITRGTPSKISVESIPSLRSGSITQGTALPQGTPTALVKGSIIRMPEDSSPERG 1200
Qy 1201 REEAKSGHVIYEGSGHILSDYDNKARECTSPRTAHEISLKRVSVEGNIKQKSM 1260
Db 1201 REEAKSGHVIYEGSGHILSDYDNKARECTSPRTAHEISLKRVSVEGNIKQKSM 1260
Qy 1261 RESPSAPLEGLICRALPRGSPHSDLKERTVLSGSIIMOGTPRATTESPDGLKYPKQIKR 1320
Db 1261 RESPSAPLEGLICRALPRGSPHSDLKERTVLSGSIIMOGTPRATTESPDGLKYPKQIKR 1320
Qy 1321 ESPPIRAFEGATKCKPYDGTITIKEMGRSITHEIPRODILTOESRKTPEVOSTREPIIG 1380
Db 1321 ESPPIRAFEGATKCKPYDGTITIKEMGRSITHEIPRODILTOESRKTPEVOSTREPIIG 1380
Qy 1381 SISOQTPPIKFDNNSSQSAIKHNKSLTGTGSKLSRGMPPLPPIVPENIKVYERCKTEDVKA 1440
Db 1381 SISOQTPPIKFDNNSSQSAIKHNKSLTGTGSKLSRGMPPLPPIVPENIKVYERCKTEDVKA 1440

BEST AVAILABLE COPY

